DOCUMENT RESUME

ED 144 402 FL 008 908

AUTHOR Stump, Thomas A.
TITLE Cloze and Dictation Tasks as Predictors of Intelligence and Achievement. Occasional Papers on Linguistics, No. 1.
INSTITUTION Southern Illinois Univ., Carbondale. Dept. of Linguistics.
PUB DATE Apr 77
NOTE 29p.; Proceedings of the International Conference on Frontiers in Language Proficiency and Dominance Testing (1st, Southern Illinois University, Carbondale, Illinois, April 21-23, 1977)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Aptitude Tests; *Cloze Procedure; Educational Testing; Elementary Secondary Education; English; Grade 4; Grade 7; *Intelligence Tests; *Language Proficiency; Language Skills; *Predictive Ability (Testing); Standardized Tests; *Standard Spoken Usage; Test Reliability; *Test Validity
IDENTIFIERS *Dictation; Iowa Tests of Basic Skills; Lorge Thorndike Intelligence Test

ABSTRACT

This study attempts to determine the extent to which performance on educational tests is dependent upon language proficiency. Three questions are involved: (1) Will the dictation and the cloze tests prove to be as successful in distinguishing degrees of proficiency among native speakers as among second language learners? (2) How well will scores on these pragmatic tests correlate with scores on the ITBS (Iowa Tests of Basic Skills), designed to measure supposedly separable language skills? (3) What is the degree of intercorrelation between scores on the two pragmatic language tests, the ITBS subtests, and both the verbal and nonverbal subtests of the Lorge-Thorndike Intelligence Test, which purports to be a series of tests of abstract intelligence? Tests were administered to a sample representative of the St. Louis school population (4th and 7th graders). The results seems to suggest that the measures designed to assess intelligence are so greatly dependent upon proficiency in standard English that they consequently offer a much more accurate assessment of the level of that proficiency than of abstract intelligence. All tests are essentially measuring the same thing, global language proficiency. (Author/CFM)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

James E. Redden, Editor

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

ED144402

Cloze and Dictation Tasks as Predictors of Intelligence and Achievement

Thomas A. Stump
Southern Illinois University

Because at least two of the three "R's" (readin' and writin') necessarily involve the ability to use language, it follows that language is the key ingredient in getting an education. By the same token, for some students language looms as the major obstacle to progress through the school system. Educational tests, for instance, are designed to measure both potential to gain from and results of education and are necessarily dependent to a large extent on language proficiency. Because "it is becoming increasingly clear that... educational failure is primarily a linguistic failure" (a review of Halliday, 1977), the present study has been designed to determine empirically the extent to which performance on two widely used standardized tests--the Lorge-Thorndike Intelligence Test and the Iowa Tests of Basic Skills--is dependent upon proficiency in the so-called "standard" variety of English.

Due to the importance of this topic, one might expect studies to have addressed it. In an article entitled "Interrelationships of Reading Ability, Listening Skill, Intelligence and Scholastic Achievement," Vineyard and Bailey (1960) tried to answer a version of the question posed here. They used 114 second semester freshmen and intercorrelated scores on three tests: The Reading section of the Cooperative English Test; the Listening test of the Sequential Tests of Educational Progress; and the American Council on Educational Psychological Examination. Likewise, they also used the students' one year grade point averages as measures of scholastic achievement. Correlations are shown in Table 1.

Table 1.
Intercorrelations between Supposed Measures of Reading, Listening, Intelligence and Achievement (N=114).

Tests	Reading	Listening	Intelligence	Achievement
1.Reading	--	.708	.758	.531
2.Listening		--	.671	.560
3.Intelligence			--	.546
4.Achievement				--

They conclude that "the three symbolic abilities [reading, listening, and intelligence] are closely intercorrelated indicating a great deal of relatedness of these traits, a high degree of communality in measurement devices, or both" (p.176).

A recent issue in linguistics is whether or not all language processing tasks are governed by what has been called global language proficiency. Spolsky offers a rather simplified version of

FL008908

what is meant by this term when he states that the "...ability to operate in a language includes a good, solid central portion (which I'll call overall proficiency)..." (1975:69). Some have argued to the contrary that language proficiency can be divided into separate linguistic components (phonology, syntax, lexicon) and skills (reading, writing, speaking, listening) in such a way that each component or skill can be taught and tested in isolation. Those who adhere to this point of view are said to have adopted a discrete point philosophy (using a term introduced by Carroll, 1961). Others claim that language proficiency is a much more unitary phenomenon whose "...components..., whatever they may be,...function more or less similarly in any language based task" (Oller and Hinofotis, in press). The latter viewpoint concerning the indivisibility of language proficiency has been referred to as an integrative philosophy.

In keeping with the linguistic philosophy behind them, discrete point tests are designed to test separately, in a "one-at-a-time" fashion a person's level of proficiency in each of the previously mentioned components and skills. Because it is assumed that language proficiency can be compartmentalized, the discrete point tester is not at all surprised if the tests show proficiency in one linguistic component or skill to be uncorrelated or only moderately correlated with proficiency in another component or skill. In fact, in discrete point language testing, it is expected by definition that low correlations will be obtained between different tests (or subtests) since each is supposedly measuring a separate linguistic component or skill. Thus, it is considered normal for a test of phonology to correlate minimally or not at all with a test of vocabulary. Likewise, if a listening test rates students as highly proficient while a reading test rates them as poor to average, the discrete point tester does not consider this as out of the ordinary.

The adoption of an integrative philosophy, however, requires that, in teaching and testing, language must be treated in a more holistic manner. This is the philosophy behind the observation that:

> What makes it [discrete point analysis] unworkable as a basis for teaching or testing language is that crucial properties of language are lost when its elements are separated. The fact is that in any system where the parts interact to produce properties and qualities that do not exist in the parts separately, the whole is greater than the sum of its parts. If the parts cannot just be shuffled together in any old order--if rather they must be put together according to certain organizational constraints--those organizational constraints themselves become crucial properties of the system which simply cannot be found in the parts separately (Oller, in press).

Thus, when a discrete point test requires a person to make a phonemic contrast, inflect a verb, or define a lexical item in isolation and without any of these organizational constraints, "certain crucial properties of language" are necessarily lost.

In other words, because discrete point tests do not require a person to deal with language in the way he normally does while watching television, conversing with friends, or reading or writing letters, and so on, they have the potential of distorting estimates of proficiency.

In place of discrete point tests, Oller has argued in favor of a special class of integrative tests, namely pragmatic language tests.

> Any proposed testing procedure which is to qualify as a pragmatic language processint task must meet...two naturalness criteria...: (1) It must require the processing of temporal sequences of elements in the language constrained by the normal meaningful relationships of such elements in discourse, and (2) it must require the performer of the task to relate the sequences of elements to extralinguistic context via...pragmatic mappings. More succinctly, pragmatic tasks require time constrained processing of the meanings coded in discourse (in press).

Because pragmatic language tests are designed to meet these three strict requirements--and as a result to significantly tap what Spolsky calls the "good, solid central portion" of overall language proficiency--all such tests are expected to intercorrelate at high levels regardless of which traditionally recognized skill (listening, speaking, reading, or writing) is specifically utilized. More importantly, because the major strength of the pragmatic test "lies in the fact that it does reflect a natural use of language" (Bondaruk, Child, and Tetrault, 1975:90), it not only has the potential to provide a much more valid assessment of language proficiency, but it is more apt to offer insights into the workings of the human mind. Indeed, a person performing a pragmatic task is not dealing with isolated fragments of language; instead, he is involved in "processing...temporal sequences of language constrained by... normal meaningful relationships" and relating them "to extralinguistic context via...pragmatic mappings" while operating under "time constrained processing of...meanings".

Consequently, it becomes evident that the general question addressed by this study (concerning the extent to which performance on educational tests is dependent upon language proficiency) involves at least three subsidiary questions: First, will the two pragmatic tests used in this study--the dictation and the cloze--prove to be as sucessful in distinguishing degrees of proficiency among native speakers as they have been when used with second language learners? Secondly, how well will scores on these pragmatic tests correlate with scores on the ITBS subtests designed to measure supposedly separable language skills? (In other words, will this study provide evidence for or against an integrative philosophy and the notion of global language proficiency?) Finally, what is the

degree of intercorrelation between scores on the two pragmatic language tests, the ITBS subtests constructed in a discrete point manner, and both the verbal and nonverbal subtests of the Lorge-Thorndike which purport to be a "series of tests of abstract intelligence" (Manual for Administration, p. 4)? By addressing the subsidiary questions in this order, we thus arrive at a more succinct rendering of this study's primary question: Is it possible that all three of these tests are measuring basically the same thing--namely, global language proficiency?

Method

In January, 1977, a research proposal was sent to and accepted by the St. Louis Board of Education. The proposal indicated that the major purpose of the research was to intercorrelate scores from an achievement test, an intelligence test, and a pragmatic language test in order to determine patterns of interrelationship and amount of variance overlap among these tests. Scores from the Lorge-Thorndike Intelligence Test (administered in October, 1976) and from the Iowa Tests of Basic Skills (given in April, 1977) were obtained at the time of the administration of the pragmatic language tests (April and May, 1977).

Subjects: The subjects used in this study were selected by the Board of Education from four different schools in the St. Louis public school system. There were 109 fourth grade students and 95 seventh grade students selected, setting the total number of subjects at 204. These two grades were chosen primarily because they were the only ones to take the Lorge-Thorndike Intelligence Test during this school year. All students in grades 3 through 8 are required to take the ITBS.

In the proposal it was requested that a sample representative of the St. Louis school population be obtained. Based upon the mean scores and the standard deviations of the ITBS and Lorge-Thorndike subtests (see tables 2 and 6 in the results and discussion section), the sample obtained appears to be nearly identical to the sample population on which these tests were standardized. However, because 85% of the 204 subjects tested were white (from a school system which is more than 70% black), it would be rather difficult to say whether or not the sample was really representative of the St. Louis public schools. (Suffice it to say that court-ordered school desegregation has had a rather chaotic effect upon the St. Louis public school system--a few of the black students who were tested had already attended four different schools in as many years.) Nonetheless, it is really not at all crucial to this study who takes the tests, since the question addressed here is whether anyone taking these tests will perform in the same way or differently on all of them.

Tests: The particular edition of the ITBS used in this study was copyrighted in 1964 by the State University of Iowa under the direction of E.F. Lindquist and A.N. Heironymus. The Iowa Tests of Basic Skills "provide for comprehensive measurement of the following fundamental areas: vocabulary, reading, the mechanics of correct writing, methods of study, and arithmetic. These skills are crucial to the educational development of the pupil. They largely determine the extent to which he can profit from later instruction" (Teacher's

Manual, p.3).

The following specific skills are tested, with the time allotted for the taking of each subtest in parentheses: Vocabulary (17 minutes); Reading Comprehension (55 minutes); Language Skills (67 minutes), composed of Spelling (12 minutes), Capitalization (15 minutes), Punctuation (20 minutes), and Usage (20 minutes); Work-Study Skills (80 minutes), comprised of Map Reading (30 minutes), Reading Graphs and Tables (20 minutes) and Knowledge and Use of Reference Materials (30 minutes); and Arithmetic Skills (60 minutes), which includes Arithmetic Concepts (30 minutes) and Arithmetic Problem Solving (30 minutes). Thus, there are actually eleven separate subtests to be taken by students from grades three through nine, with total testing time amounting to nearly five hours.

In completing the vocabulary subtest, a student is required to select from four choices the word or words most closely related in meaning to another word embodied in a short phrase. For example:
Say something humorous.
1) comical 2) intelligent 3) full of wisdom 4) interesting
(from the 4th grade section)
Each grade level is required to complete between 35 and 50 such exercises.

The reading comprehension subtest requires that a student read up to 10 passages of varying lengths while answering several multiple-choice questions on each passage. For instance, the final passage read by the seventh grade students was an approximately 600-word biography of General George Custer. The selection was followed by 21 comprehension questions, nearly all of which began with words like who, which, what, where, how old, how many, etc.

All four sections of the language skills subtest follow the same format: The student is required to recognize the presence or absence of errors. The directions for each of these sections includes the clause, "You are to look for mistakes in..." Examples:
(Spelling)
1) original 2) patients 3) chemestry 4) responsible 5) (No mistakes)

(7th Grade)

(Capitalization)
1) The twins were ashamed.
2) They didn't know where
3) New zealand was on the map.
4) (No mistakes)

(4th Grade)

(Punctuation)
1) The goats had eaten
2) almost every plant in the
3) garden, Nothing was left.
4) (No mistakes)

(4th Grade)

(Usage)
1) This brick walk isn't a very good place for roller skating.
2) The fisherman was grouchy because he didn't catch nothing.
3) There are hardly any flowers left after last night's heavy frost.
4) (No mistakes)

(7th Grade)

It should be noted that none of these subtests meet any of the requirements of a pragmatic language task. Instead, all have been constructed in a strictly discrete point fashion. By testing "the mechanics of correct writing," the first three sections of the language skills subtest have in fact taken the discrete point approach to its logical extreme. In addition, the usage section appears to be quite removed from the actual use of language when one considers that all of its exercises require a student to label certain so-called "non-standard" usages (common to many regions and social strata of the United States) as "incorrect". Such insensitivity on the part of this section to the dialectal differences which exist in this country must undoubtedly create bewilderment in the minds of many children.

The students in the St. Louis school system were not required to take the work-study skills subtest this year. It will suffice here to say that a student taking this subtest must answer numerous multiple-choice questions based on an array of highway and topographical maps, mathematically-oriented tables, charts, and graphs, and reference materials such as the dictionary.

Two examples will readily show what is demanded of the student taking the Arithmetic Skills subtest:

(Arithmetic Concepts)
Which of these examples does <u>not</u> have the same answer as the other three?

1)	2)	3)	4)
.394	3.94	394	39.4
<u>x27</u>	<u>x.27</u>	<u>x.0027</u>	<u>x.027</u>

(7th Grade)

(Arithmetic Problem Solving)
There were 29 fourth grade pupils in Washington School at the beginning of the year. During the school year, 4 new pupils joined the class. None moved away. How many pupils were in the fourth grade then?
1) 23 2) 29 3) 32 4) (not given)

(4th Grade)

Due to the highly verbal nature of math problems such as these, it would seem logical to anticipate at least a moderate degree of correlation between the arithmetic skills subtest and the other language-based subtests of the ITBS.

When one examines simultaneously the contents of both the Lorge-Thorndike Intelligence Test and the Iowa Tests of Basic Skills, one is struck by the high degree of similarily between these two measuring devices. The latest edition of the Lorge-Thorndike (prepared at the Teacher's College, Columbia University, by Irving

Lorge, Robert L. Thorndike, and Elizabeth Hagen) bears the same copyright date (1964) as the ITBS; and likewise both are published by the same company (Houghton Mifflin).

The Lorge-Thorndike is composed of two subtests--one verbal, the other (purportedly) nonverbal.

> The tests are based on the premise that...[because]... most abstract ideas...are expressed in verbal symbols,...verbal symbols are the appropriate medium for the testing of abstract intelligence. Nevertheless, because...printed words may constitute an inadequate basis for appraising an individual's abilities,...a parallel set of nonverbal tests is provided to accompany the basic verbal series (Manual for Administration, p. 4).

Thus, because of its admittedly heavy verbal weighting, one would expect a rather high degree of correlation between the verbal subtest and various other language-based tasks (as predicted by Vineyard and Bailey). In addition, because the nonverbal portion has been included not to measure something distinct from the verbal section, but instead to accompany and be parallel to it, one would consequently expect also a substantial degree of intercorrelation between both of these subtests and various kinds of language-oriented tasks. Evidence in support of such a prediction is also provided by a further quotation from the Manual:

> It is recommended that the administrator... practice a natural and meaningful speech pattern when reading the directions to the pupils. If the directions are read without proper stress, some pupils will fail to understand what they are supposed to do, and their lack of understanding of the testing procedure will be reflected in unreliable test scores (page 6).

Reading the directions with a "natural and meaningful speech pattern" including "proper stress," however, will have little effect on the performance of a student whose level of language proficiency simply prevents his from fully comprehending what he is being asked to do. As a result, the test scores--while being invalid and unreliable in the sense that they do not measure the construct for which the test was designed--may prove to represent a valid and reliable assessment of language proficiency.

A student is given 35 minutes to complete the five sections of the verbal subtest (7 minutes for each one). Its format is nearly identical to that of the ITBS. First, there is a section called Vocabulary...

triangular-
a) geodetic b) surveying c) surveyed d) three sided
e) equilateral

(4th Grade)

followed by Sentence Completion:

It is much easier to be critical than to be _____.
a) false b) spurious c) free d) correct e) disparaging
(7th Grade)

The Arithmetic Reasoning section contains math problems so similar to those on the ITBS that one would be hard pressed to tell them apart:

The cost of a round trip ticket between cities A and B is $5.45. A one-way ticket costs $2.90. How much money would a family save by buying three round trip tickets instead of six one-way tickets?
a) $1.15 b) $ 1.95 c) $3.10 d) $.99 e) None of these
(4th Grade)

In the Verbal Classification section, one is required to first discover the way in which four words are alike and then select among five choices the word which belongs with the group:

gift, bequest, legacy, donation
a) subscription b) just claim c) lease d) contribution
e) collection (4th Grade)

Finally, there is the Verbal Analogy section:

fertility→fruition: sterility→_____
a) destruction b) disruption c) dissolution d) extinction
e) subversion

In all three sections of the nonverbal subtest (Pictorial Classification, Pictorial Analogy, and Numerical Relationships--each requiring 9 minutes of work time), a student is directed verbally to select from five choices the shape, picture, or number which belongs with a group of other such shapes, etc. It has been seen, however, that the dependence of this subtest upon verbal instructions tends to place its supposedly nonverbal character in jeopardy. (For a more detailed account of the content validity of these and other standardized tests, one is referred to Gunnarsson, in press.)

Construction, administration, and scoring of pragmatic tests: One important consideration in constructing pragmatic tasks is that the passages of contextually connected discourse from which the tasks are fashioned be of interest to the population tested. Every effort was made to select (from books, magazines, etc.) and create passages with which the inner-city student could readily identify. For instance, one selection involved a young boy who enjoyed riding his new bike around the neighborhood.

Passages were also chosen and created according to some general notions of lexical and syntactic complexity: Certain

lexical items (e.g. skeptic, zodiac, ancestor) and syntactic features (subordinations and specific verb tenses) were considered more complex and consequently more difficult than others. Based upon such notions, eight passages (four to be fashioned into dictation tasks, four into cloze tasks) were selected according to four degrees of difficulty: easy, medium, somewhat hard, quite hard.

During the construction of the tasks themselves, additional steps were taken to ensure an ordered increase in the level of difficulty. With respect to dictation, this was done as the tasks were tape-recorded. Each passage was read three times at a normal conversational speed: The first time, students were asked to listen for the main idea. The second time, they were required to write down exactly what they heard. (This was done during pauses appropriately spaced throughout the passage. Pauses were always long enough to allow adequate time for slow writers, since dictation is not intended as a test of writing speed.) The passage was then read a third time while the students checked what they had written.

Although these steps were followed in exactly the same manner for all four tasks, there was one crucial difference from task to task: During the second reading of the passage, the length of each utterance to be recalled and written was highly controlled. Thus, Dictation B required the students to recall and write more words (an average of 5.0 per utterance) than Dictation A (3.0), Dictation C more words (9.5) than Dictation B, and so on. By increasing or decreasing the average length of utterance (ALU) in this way, one can greatly affect overall performance regardless of the lexical or syntactic complexity of the passage.

In constructing cloze tasks, one can also alter the level of difficulty despite passage content. Even when a fixed-ratio method of deletion is adopted (as was the case here whereby every seventh word was omitted and replaced with a blank), one can still choose the types of words to be deleted by simply editing or rearranging the passage itself. For instance, while the words selected for deletion in Cloze A were primarily rather simple nouns, pronouns, main verbs, and articles, the omissions in Cloze B included several prepositions, auxiliary verbs, and subordinate conjunctions, those in Cloze C some rather uncommon adverbs and adjectives, and so on.

Before the actual testing was begun in St. Louis, a highly informal test situation was created by using several youngsters from the Carbondale, Illinois, public school system. This was done in order to make certain that the difficulty levels selected were appropriate for the population to be tested. The results were indeed quite promising: The tests proved to be neither too easy nor too difficult, and in addition they even ranked the youngsters in the same order as did their teacher.

It was decided that the pragmatic tests would be administered to no more than forty students at a time, thus making it easier for the administrator to check each student's progress at various intervals during the testing period. Once each student had received a test booklet (a copy of which is provided in the appendix), the directions for the dictation subtest were read orally as the students followed along. Additional comments were interjected in order to clarify such points as the treatment of punctuation. Then, while

a sample dictation of 15 words was being played, the administrator made sure that all students fully understood exactly what they were being asked to do.

This having been done, the actual testing commenced. Approximately 15 minutes was required for the playing of the dictation tasks which ranged in length from 55 to 86 words. The cloze subtest (comprised of four, 20 item tasks) was administered in exactly the same manner, except that students were given 30 minutes to complete the tasks. Only one student ever failed to finish because of this time restriction.

Although eight separate tasks were constructed, each grade was required to complete only six: The 4th grade, dictation and cloze tasks A,B, and C; the 7th grade, tasks B,C, and D. Such a procedure prevented students in the two grades from performing tasks which were too easy or too difficult, while still allowing a comparison between the grades on four separate tasks (dictation and cloze tasks B and C).

In scoring the dictation tasks, each word in a passage was worth one point. Consequently, the maximum scores on Dictations A through D were 55, 66, 76, and 86 points respectively, setting the total possible dictation score for the 4th grade at 197, the 7th grade at 228. One point was deducted for every word deleted from the passage, and likewise for every word inserted unnecessarily. In addition, each word which was distorted phonologically or morphologically ("He road his bike" instead of "He rode his bike") also resulted in the loss of a point. (Because research has indicated a very weak relationship between spelling and ability to sucessfully perform various kinds of pragmatic tasks, it was decided that spelling errors would not be counted: "naberhood" in place of "neighborhood" was scored as correct.) Also, if a word or phrase were placed in a different location within the sentence ("There is near our school a big tree..." instead of "There is a big tree near our school..."), one point was deducted.

Each blank in the cloze tasks was likewise given a value of one point, setting the maximum score for each of the four tasks at 20, the total possible cloze score for each grade at 60. Two commonly used methods of scoring cloze tasks were considered: The exact word method, whereby one is required to reproduce the very word which was deleted; and the contextually appropriate method, whereby any word is accepted as long as it conforms to all the constraints imposed by the surrounding context. Although research with second language learners has shown consistently high correlations between these two scoring procedures (students were ranked in nearly identical orders regardless of the method used), the contextually appropriate method was preferred here for use with native speakers.

Consequently, in addition to the exact word, any other choice which was not in violation of either local (structural) or long range (involved with the overall meaning) constraints was also scored as correct. Surprisingly (perhaps), in over 80% of the cases the blank was filled either with the exact word deleted or with a word in obvious violation of some local constraint ("I've have

tossing coins for five years now" instead of "I've _been_ tossing..."). In most of the remaining cases, the choice was either clearly appropriate ("So they _went_ over to Bill's house..." in place of "...they _walked_ over...") or else in violation of a long range constraint ("I only have three _friends_ and you don't have any money _at_ all..." instead of "...three _dollars_ ..." or "three _cents_ ..."). In the few cases where some degree of judgment was involved, another native speaker was called upon for a second opinion. The student was always given the benefit of the doubt.

Results and Discussion

The findings of this research will be presented in two stages: First, the three subsidiary questions will be addressed with respect to each grade separately; then, there will follow a concluding discussion of the primary issue at stake in this study.

In terms of the 4th grade, it was found that the cloze and dictation tasks produced a great deal of variance (the dispersion of scores about the mean, or average score) among the 109 students tested. Total dictation scores ranged from a low of 85 to a high of 197 (the only perfect score), producing a mean score of 166 and a standard deviation of 22. Total cloze scores ranged from 18 to a perfect score of 60 (obtained by the very person who performed the feat on the dictation tasks), resulting in a mean score of 44, a standard deviation of 8. The mean scores, standard deviations, and cases of the six individual pragmatic tasks, as well as the two subtests of the Lorge-Thorndike and the three subtests of the ITBS, are shown in Table 2.

Table 2.
Means, Standard Deviations, and Cases of the Six Subtests of Dictation and Cloze, the Two Subtests of the Lorge-Thorndike Intelligence test, and the Three Subtests of the Iowa Tests of Basic Skills (4th Grade).

Subtest	Mean	Standard Dev.	Cases
1 Dict. A	48.9	6.5	109
2 Dict. B	59.4	7.6	109
3 Dict. C	57.5	11.2	109
4 Cloze A	16.4	2.5	109
5 Cloze B	14.7	3.2	109
6 Cloze C	13.1	3.7	109
7 LT-Non V.	102.0	14.9	95
8 LT-Verbal	102.5	13.6	95
9 ITBS-V/R	4.9	1.1	108
10 ITBS-Lg.	5.5	1.0	107
11 ITBS-Ar.	5.0	1.0	107

"LT-Non V." and "LT-Verbal" are themselves means of scores from the various sections of the two respective subtests. "ITBS-V/R" represents the mean score obtained when the Vocabulary and

Reading Comprehension subtests are combined. (In hindsight, the researcher regrets that the scores from these two subtests were not treated separately, for this could have proved advantageous to the study's address of the integrative vs. discrete point controversy.) "ITBS-Lg." signifies the mean scores from the four sections of the Language Skills subtest, "ITBS-Ar." the mean of scores from the two sections of the Arithmetic Skills subtest.

The ordered decrease in mean scores and consequent increase in standard deviations of the pragmatic tasks bear out quite well the predicted rise in degree of difficulty from task to task. As mentioned previously, the mean scores and standard deviations of the Lorge-Thorndike verbal and nonverbal subtests are shown to be nearly identical to those considered "normal" for intelligence tests in general: 100 and 15, respectively. As the "cases" column indicates, 14 of the 4th graders tested had not taken the Lorge-Thorndike. In addition, it should be noted that scores on the ITBS are expressed in years. For instance, the mean of the language skills subtest is read "fifth year, fifth month", its standard deviation "one year".

It is evident from Table 2 that the close and dictation tasks, the Lorge-Thorndike, and the ITBS have all produced substantial amounts of variance among these 4th graders, and thus there is good indication that each one may in fact be ranking students according to the construct for which it was designed (global language proficiency, abstract intelligence, and basic skills or achievement, respectively). A test, however, is not usually regarded as a valid measuring device until its designer offers evidence that the variance it produces correlates rather highly with the variance produced by other tests designed to measure the same construct, and thus establishes its concurrent validity. Correlation, in fact, expresses the degree of variance overlap between two tests (the percentage of variance which they share in common), for by simply squaring the correlation coefficient between two tests (.80, for example), one arrives at the percentage of variance overlap (.64).

Consequently, as designer of the pragmatic tests used in this study, the author offers as evidence of their concurrent validity the fact that cloze and dictation tasks (identical to those used here) consistently correlate quite highly not only with various other pragmatic tasks designed to tap global proficiency, but also with discrete point tests which purportedly measure vocabulary, grammar, composition, etc. (Oller and Streiff, 1975). Such evidence has given strong impetus to the integrative point of view. Likewise, designers of the Lorge-Thorndike base the validity of their test primarily upon the fact that it correlates rather highly (.69's, .70's, and low .80's) with other purported group tests of intelligence.

Nowhere in the Teacher's Manual of the ITBS, however, is there any indication that its subtests have been correlated with other tests designed to measure the same things (vocabulary, reading comprehension, etc.). Instead, it appears that the ITBS has been tacitly granted validity by virtue of its close relationship with

the Lorge-Thorndike, for in addition to their striking similarities in terms of format, content, etc., one discovers evidence of an even more intimate association between the two tests:

> The relationship between intelligence test scores and achievement leads one to expect that students who score high on an intelligence test are likely to stand high in achievement and those who score low are likely to be low in achievement. Since...the Lorge-Thorndike Tests and Iowa Tests of Basic Skills have been standardized together, one can compare the individual's percentile rank on the two tests and plot the results of both tests on a single profile (Lorge-Thorndike Manual, p.38).

Consequently, what degree of correlation should one anticipate between the ITBS and the Lorge-Thorndike?

> Although one would expect that a student who scored at the 99th percentile on the Lorge-Thorndike Test would score high on the Iowa Tests of Basic Skills, one should not expect a one-to-one correspondence on the two tests; that is, one should not expect the student to fall exactly at the 99th percentile on the Iowa Tests of Basic Skills..., since both scores contain errors of measurement (p.38).

Because the tests' designers predict not just a substantial, but indeed a nearly perfect correlation between the two measuring devices, one must logically ask how either test could claim to measure something the other does not; for if the correlation between the ITBS (purportedly a test of achievement) and the Lorge-Thorndike (a supposed measure of abstract intelligence) proves to be as high or higher than the correlation between the Lorge-Thorndike and other purported measures of abstract intelligence, wouldn't this indicate a rather unscientific utilization of concurrent validity? In other words, if Tests A and B intercorrelate at, say, .70, and Tests B and C at .75, how can one justifiably claim that tests A and B are measuring the same thing, while at the same time claiming that Test C is measuring something different from Test B?

Table 3 shows the intercorrelations between the subtests of the Lorge-Thorndike and the ITBS, the ITBS composite (the mean of scores from its subtests), the dictation and cloze totals, "Language" (dictation total plus cloze total), and the Lorge-Thorndike total for the 4th grade (N=95-109).

Table 3.

Intercorrelations between the Two Subtests of the Lorge-Thorndike Intelligence Test, the Three Subtests of the Iowa Tests of Basic Skills, the ITBS Composite Score, the Dictation and Cloze Totals, Language (=Dictation Total plus Cloze Total), and the Lorge-Thorndike Total (4th Grade).

Subtest/Total	1	2	3	4	5	6	7	8	9	10
1 LT-Non Ver.	---	.67	.59	.44	.61	.62	.39	.53	.47	.92
2 LT-Verbal		---	.70	.55	.59	.70	.63	.66	.70	.91
3 ITBS-Voc/R			---	.60	.71	.88	.51	.70	.62	.70
4 ITBS-Lang.				---	.73	.87	.29	.64	.42	.54
5 ITBS-Arith					---	.91	.41	.56	.49	.66
6 ITBS-Compos.						---	.45	.71	.57	.72
7 Dict. Total							---	.57	.97	.55
8 Cloze Total								---	.75	.65
9 Language									---	.63
10 LT-Total										---

Correlations range from .29 between the dictation total and the ITBS language skills subtest to .73 between the latter and the ITBS arithmetic skills subtest (discounting correlations between total or composite scores and the individual subtests which comprise those scores, found in columns 6, 9, and 10). In addition to the ITBS language skills subtest, the dictation total also correlates rather poorly with the Lorge-Thorndike non-verbal (.39) and ITBS arithmetic skills (.41) subtests; and as a result, "Language" likewise correlates somewhat weakly with those three subtests.

With the exception of correlations between the Lorge-Thorndike nonverbal and ITBS language skills subtests (.44), and between the dictation total and the ITBS composite (.45), the remaining correlations fall between the low .50's and low .70's, with numerous correlations in the mid to high .60's and low .70's. The cloze total is found to correlate rather highly with the ITBS language skills subtest (.64), the Lorge-Thorndike verbal subtest (.66), the ITBS vocabulary/reading subtest (.70), and the ITBS composite (.71). In addition to the cloze total, the Lorge-Thorndike verbal also correlates rather well with the dictation total (.63), the ITBS vocabulary/reading subtest (.70), the ITBS composite (.70), and "Language" (.70). The Lorge-Thorndike total correlates rather strongly with the cloze total (.65), the ITBS arithmetic skills subtest (.66), the ITBS vocabulary/reading subtest (.70), and the ITBS composite (.72). In addition, the ITBS vocabulary/reading and arithmetic skills subtests intercorrelate at .71, the two subtests of the Lorge-Thorndike at .67, and the cloze and dictation

totals at a somewhat lower .57.

From the correlation matrix in Table 3, one finds evidence of substantial interrelationship between all the subtests and totals, excepting somewhat the dictation total. In order to quantify the actual strength of this or any other interrelationship within the correlation matrix--and thus to better explain the overall patterns of intercorrelation--one applies a statistical process known as factor analysis. In performing such an analysis, the computer seeks out general patterns of interrelationship (called factors) which will account for specific amounts of overlapping variance within the correlation matrix. Factor 1 is the interrelationship which accounts for the greatest amount of overlapping variance, Factor 2 the next greatest amount, and so on. Simultaneously, the computer calculates the extent to which each variable in the matrix loads on (correlates with) each particular factor. In other words, it determines how substantial a role each variable plays in the specific interrelationship or factor in question. (One should note that although dictation, cloze, and other total scores were presented in Table 3 as a matter of convenience and efficiency, the factor analysis itself was actually performed on the eleven original variables: the six cloze and dictation tasks, the two Lorge-Thorndike subtests, and the three ITBS subtests.) The factor analysis is completed when all the variance in the correlation matrix has been accounted for, or exhausted.

One might more readily appreciate the process of factor analysis by attempting to predict the first few factors in the 4th grade correlation matrix. Factor 1 should be expected to represent the moderate to strong interrelationship which exists between all the subtests (and totals), and it will most probably be found to account for quite a substantial amount of the total overlapping variance in the matrix. One should anticipate the heaviest loadings on this factor to come from those variables (subtests) which share the highest degrees of intercorrelation with the largest number of other variables. These should include the Lorge-Thorndike verbal subtest, the cloze tasks, the ITBS vocabulary/reading subtest, and the ITBS arithmetic skills subtest, with lower weightings (loadings) coming from the other variables, the lowest from the dictation tasks. Because of their low to moderate degrees of interrelationship with other variables, one would expect the dictation tasks to load to a certain extent on a separate factor (probably the second one). Likewise, since the Lorge-Thorndike non-verbal subtest shares some rather low degrees of intercorrelation, one might expect it to weigh on the third or fourth factor to a somewhat limited extent.

Table 4 shows the eigen values of and percents of variance exhausted by each of eleven factors. (With respect to the meaning of these eigen values, it will suffice to point out the intimate, almost one-to-one association between them and their corresponding percents of variance.

Table 4.

Eigen Values and Percents of Variance for Eleven Factors (4th Grade)

Factor	Eigen Value	Pct. of Var.
1	6.0	54.2
2	1.4	12.6
3	.9	8.0
4	.6	5.5
5	.5	4.5
6	.4	4.0
7	.3	3.1
8	.3	2.9
9	.2	2.2
10	.2	1.7
11	.1	1.3

As expected, factor 1 is shown to account for quite a large amount of variance overlap (54%), and with the addition of factor 2, nearly 70% of the total variance in the correlation matrix is exhausted. Individual loadings of the eleven variables on these two primary factors are shown in Table 5.

Table 5.
Factor Matrix Using Principal Factor (No Iterations) for the Six Subtests of Dictation and Cloze, the Two Subtests of the Lorge-Thorndike Intelligence Test, and the Three Subtests of the Iowa Tests of Basic Skills (4th Grade).

Subtest	Factor 1	Factor 2
1 Dict. A	.60	.64
2 Dict. B	.64	.53
3 Dict. C	.71	.49
4 Cloze A	.66	-.12
5 Cloze B	.76	-.15
6 Cloze C	.82	-.02
7 LT-Non Ver.	.71	-.19
8 LT-Verbal	.85	.04
9 ITBS-Voc/R.	.84	-.18
10 ITBS-Lg.	.71	-.48
11 ITBS-Ar.	.76	-.34

Weightings on factor 1 range from .60 for Dictation A to .85 for the Lorge-Thorndike verbal subtest. As predicted, this factor also receives heavy weightings from the ITBS vocabulary/reading subtest (.84) and Cloze C (.82), as well as from the ITBS arithmetic skills subtest and Cloze B (both .76). Dictation C, the Lorge-Thorndike nonverbal subtest, and the ITBS language skills subtest all load on factor 1 at .71. For each individual variable in the factor matrix, one may calculate the percentage of variance accounted for by this or

any other factor simply by squaring the weighting of the variable on the factor in question, since that weighting actually represents the correlation between the variable and factor. If one then adds the resulting percentages for all the variables and divides by the total number of variables, one arrives at the total variance accounted for by the particular factor.

Although the second factor is clearly a "dictation" factor (essentially all of its positive loading is exhausted by the dictation tasks), more of the total dictation variance is still accounted for by factor 1 (42%) than by factor 2 (31%). This indicates that despite some rather weak intercorrelations, the dictation tasks are undeniably measuring something more in common with than in divergence from what the other subtests are tapping. One may note the role of difficulty level in determining this outcome, for as the degree of difficulty increased from task A to B to C, so was there a simultaneous increase in the amount of variance shared with other subtests. Subsequently, while Dictation C weighs the heaviest of the three tasks on factor 1 (.71) and the least heavy on factor 2 (.49), the inverse is true for Dictation A. Likewise, the more difficult the cloze task, the greater its amount of overlapping (shared) variance and the heavier its subsequent loading on factor 1.

Thus, considering the large amount of variance they have produced among these 4th graders and the fact that this variance overlaps to a moderate to substantial degree with that produced by all of the subtests of the Lorge-Thorndike and ITBS, one finds quite promising evidence for the capability of cloze and dictation tasks to distinguish degrees of proficiency among native speakers. In fact, the results here seem to indicate a potential for pragmatic tasks such as these to reveal considerably more variation in terms of native language proficiency than was ever imagined to exist.

With respect to the discrete point versus integrative controversy, one discovers support for both points of view. On the one hand, the substantial degree of variance overlap (or covariance) between the cloze tasks and both the ITBS language skills and vocabulary/reading subtests, as well as between the latter two subtests themselves, lends tenability to the integrative approach. (As mentioned previously, a separate treatment of scores from the vocabulary and reading comprehension subtests would have unquestionably added an extra dimension to the study's address of this important question; the oversight is indeed regrettable.) On the other hand, since the dictation tasks are obviously tapping something which these subtests are not, one also finds support for a separate skills approach, possibly interpreted as reading versus listening skill.

The author speculates, however, that despite the precautions taken to ensure against it, writing speed may have played a key role in this divergence; for during the administration of the dictation tasks, it was observed that many of the 4th graders were frequently hard put to finish writing what they had heard before the onset of the next utterance. Consequently, it seems realistic to entertain the possibility of factor 2's representing "speed of writing" as much as or even more than a truly separate linguistic skill ("listening ability"), and perhaps we should postpone any conclusions until the

presentation of results for the seventh graders (who were much less intimidated by the entire dictation-taking process).

In terms of the third subsidiary question, one discovers that the Lorge-Thorndike verbal subtest intercorrelates quite substantially (.55 to .70) with all the other subtests in the correlation matrix. As a result, it exhausts 73% of its total variance on factor 1, the unifying or unitary factor (to use the term introduced by Oller and Hinofotis, in press), leaving only 27% to be accounted for by the following 10 factors. The Lorge-Thorndike nonverbal subtest intercorrelates with other subtests at lower levels (.39 to .67), but still exhausts 50% of its total variance on the unitary factor. When this is added to the 4% accounted for by factor 2, only 46% remains to be exhausted over the next 9 factors; and even if a large percentage of this remaining variance (say 35% were accounted for by one factor (as predicted), the fact would still remain that the major portion of its variance is exhausted by factor 1.

Thus, with regard to the 4th grade, we are presented with strong evidence that all the subtests are weighing substantially on (or tapping) an underlying unitary factor, although the possibility must be entertained that the dicataion tasks are measuring something in addition to this factor. Indeed, one discovers that several of the correlations between subtests measuring purportedly diverse things (abstract intelligence as opposed to vocabulary/reading) are as high or higher than those used to establish the concurrent validity of the Lorge-Thorndike itself. The crucial question: What is this factor which unites all of these measures? Before attempting an answer, let us first examine the results for the seventh grade.

In terms of variance, one finds that the pragmatic tasks have produced an almost identical amount of variation among the 7th grade students as they have among their 4th grade counterparts. Total dictation scores ranged from 109 to a perfect 228, total cloze scores from 19 to 59, resulting in mean scores and standard deviations of 189 and 25, respectively, for the dictation tasks, 45 and 9 for the cloze tasks. Means, standard deviations, and cases for all eleven individual subtests are presented in Table 6.

Table 6.

Means, Standard Deviations, and Cases of the Six Subtests of Dictation and Cloze, the Two Subtests of the Lorge-Thorndike Intelligence Test, and the Three Subtests of the Iowa Tests of Basic Skills (7th Grade).

Subtest	Mean	Standard Dev.	Cases
1 Dict. B	64.3	2.8	95
2 Dict. C	68.5	5.9	95
3 Dict. D	56.0	17.7	95
4 Cloze B	16.9	2.5	95
5 Cloze C	15.7	3.2	95
6 Cloze D	11.9	4.1	95
7 LT-Non V.	101.9	14.3	49
8 LT-Verbal	99.6	13.3	90
9 ITBS-V/R	7.2	1.5	93
10 ITBS-Lg.	8.1	1.5	90
11 ITBS-Ar.	7.3	2.2	91

Once again, we discover that the ordered decrease in mean scores and subsequent increase in standard deviations of the pragmatic tasks bear out perfectly the anticipated increase in diffuculty level from task to task. As predicted, dictation and cloze tasks D proved to be quite difficult, with mean scores of 56 (86 maximum) and 11.9, respectively, standard deviations of 17.7 and 4.1. A comparison of the two grades on dictation and cloze tasks B and C reveals that (as would be expected) the 7th graders performed considerably better: Whereas mean scores and standard deviations for the 4th grade were 59.4 (66 max.) and 7.6, respectively, for Dictation B, 57.5 (76 max.) and 11.2 for Dictation C, the same statistics for the 7th grade were found to be 64.3 and 2.8, 68.5 and 5.9. Similar results can be observed for the two overlapping cloze tasks.

Table 6 also shows that the means and standard deviations of both Lorge-Thorndike subtests again concentrate within two points of 100 and 15, respectively. One may note that due to a mechanical (grading) difficulty, the nonverbal scores of over 40 students at one school had not yet been received from the testing service. Thus, as the "cases" column indicates, the nonverbal N is only 49, as opposed to 90-95 for the other 10 subtests. With respect to the ITBS subtests, one finds that--like the 4th graders--the students in the 7th grade performed highest on the language skills subtest, as shown by its mean of 8.1 (eighth year, first month). The greatest amount of variance was produced by the arithmetic skills subtest which has a standard deviation of slightly over 2 years.

Despite the striking similarities in terms of variance, Table 7 reveals some equally astonsihing differences between the two grades with regard to correlation.

Table 7.

Intercorrelations between the Two Subtests of the Lorge-Thorndike Intelligence Test, the Three Subtests of the Iowa Tests of Basic Skills, the ITBS Composite Score, the Dictation and Cloze Totals, Language (= Dictation Total plus Cloze Total), and the Lorge-Thorndike Total (7th Grade).

Subtest/Total	1	2	3	4	5	6	7	8	9	10
1 LT-Non Ver.	--	.76	.66	.72	.69	.77	.70	.64	.71	.94
2 LT-Verbal		--	.73	.70	.30	.79	.70	.67	.72	.93
3 ITBS-Voc/r.			--	.69	.36	.92	.70	.75	.75	.78
4 ITBS-Lang.				--	.38	.88	.70	.64	.71	.81
5 ITBS-Arith.					--	.42	.30	.32	.32	.75
6 ITBS-Compos.						--	.76	.78	.81	.87
7 Dict. Total							--	.77	.98	.80
8 Cloze Total								--	.87	.75
9 Language									--	.82
10 LT-Total										--

Upon close examination of the matrix, one discovers a very sharp dichotomy between the ITBS arithmetic skills subtest and most of the other subtests, for it intercorrelates quite weakly (.30 to .38) with all except the Lorge-Thorndike nonverbal (with which it correlates at a substantial .69. The remaining intercorrelations in the matrix,

however, are all quite substantial, with magnitudes ranging from .64 to .87 (again discounting intercorrelations between totals and their respective subtests). The dictation total, which showed only weak to moderate intercorrelations for the 4th grade, is now ween to correlate with other subtests at a consistently high level: .70 with both subtests of the Lorge-Thorndike, as well as with the vocabulary/reading and language skills subtests of the ITBS. Intercorrelations between the cloze total and these four subtests range from .64 to .75, and the two pragmatic subtests are found to correlate at .77--just one point above the coefficient between the two Lorge-Thorndike subtests (.76). The ITBS vocabulary/reading and language skills subtests correlate with each other at .69, with the Lorge-Thorndike verbal subtest at .73 and .70, respectively, and with the nonverbal at .66 and .72. Total scores show higher intercorrelations: "Language" with the ITBS composite at .81 and with the Lorge-Thorndike total at .82; while the latter two totals are shown to correlate at a rather amazing .87.

Table 8 shows the unitary factor to account for nearly 62% of the total variance in the matrix, and with the addition of factor 2 (predictably an "arithmetic skills" factor), only 28% of the variance remains to be exhausted by the following 9 factors.

Table 8.

Eigen Values and Percents of Variance for Eleven Factors (7th Grade)

Factor	Eigen Value	Pct. of Var.
1	6.8	61.5
2	1.1	10.2
3	.7	6.3
4	.7	6.1
5	.5	4.3
6	.4	3.3
7	.3	2.8
8	.2	1.9
9	.2	1.7
10	.1	1.2
11	.1	.7

Loadings of the eleven subtests on the two primary factors are presented in Table 9. The dichotomy in the correlation matrix is just as clear-cut in the factor matrix: Whereas the ITBS arithmetic skills subtest loads heavily (exhausting 61% of its variance) on factor 2, all other subtests exhaust the major portion of their variance on factor 1. In fact, with the exception of the two easiest pragmatic tasks (Dictation and Cloze tasks B), no other subtest exhausts less than 64% of its variance on the unitary factor. Among the highest loadings are Dictation D at .87 (76% of variance), Cloze D and the Lorge-Thorndike nonverbal subtest at .85 (73% of variance), and the Lorge-Thorndike verbal and the ITBS vocabulary/reading subtests at .83 (69% of variance). Besides the ITBS arithmetic skills, the only other subtest loading to any real degree (positively) on factor 2 is the Lorge-Thorndike nonverbal (.42, or 18% of variance), reflecting the substantial intercorrelation

between these two subtests. Consequently, how should one interpret this second factor? More importantly, (and the primary issue at stake in this study), what is the most realistic interpretation of factor 1, the unitary factor? In answering these questions, it is first necessary to assess the results which we have found.

Table 9

Factor Matrix Using Principal Factor (No Iterations) for the Six Subtests of Dictation and Cloze, the Two Subtests of the Lorge-Thorndike Intelligence Test, and the Three Subtests of the Iowa Tests of Basic Skills (7th Grade).

Subtest	Factor 1	Factor 2
1 Dict. B	.71	-.37
2 Dict. C	.81	-.31
3 Dict. D	.87	-.11
4 Cloze B	.71	-.07
5 Cloze C	.80	-.18
6 Cloze D	.85	-.15
7 LT-Non Ver.	.85	.42
8 LT-Verbal	.83	.05
9 ITBS-Voc/R.	.83	.07
10 ITBS-Lg.	.81	.14
11 ITBS-Ar.	.47	.78

The cloze and dictation tasks used in this study have produced a great deal of variance among both 4th and 7th graders; and because this variance has been shown to overlap substantially with that produced by two widely used standardized tests, we find strong evidence for the capacity of these tasks to differentiate degrees of native proficiency. Such evidence comes at a time when the field of linguistics is embroiled in a fundamental dispute over the very nature of language proficiency. To briefly recapitulate: Pragmatic tasks like the cloze and dictation have been developed by those who view language as holistic and inseparable, to be taught and tested in meaningful (contextualized) situations (the integrative approach); tests like the Iowa Tests of Basic Skills, on the other hand, emanate from a belief in the divisibility of language proficiency, held by those who advocate the teaching and testing of language in a piecemeal fashion, where portions of language and supposedly separate linguistic skills are treated in isolation from each other (the discrete point approach).

For the 4th grade, we discovered substantial intercorrelations between the ITBS language skills subtest, the vocabulary/reading subtest, and the cloze tasks, providing evidence for the integrative approach. At the same time, however, weak to moderate degrees of correlation between the subtests and dictation lent support to the separate skills approach (although it was speculated that a factor other than listening skill was playing a considerable role in this

divergence). No such evidence of separability emerged for the 7th grade: All four subtests were found to intercorrelate at a consistently high level, with magnitudes ranging from .64 to .77 (the correlation between the two pragmatic subtests). Considering the fact that for the 7th grade, dictation showed no intercorrelation weaker than .70--combined with the reservations already held about positing a separate listening factor for the 4th grade--the researcher concludes that the earlier divergence on the part of the dictation tasks was most likely due to a misjudgement in their construction (which resulted in the tapping of writing speed or some other irrelevant factor), not to the existence of a separate linguistic skill.

Thus, we find that tests designed to measure vocabulary/reading and language skills (punctuation, capitalization, spelling, and usage) intercorrelate quite substantially not only with each other, but with pragmatic tasks designed to tap overall or global language proficiency. Such a finding indicates that all of these measures are tapping basically the same construct--global proficiency--and this lends strong support to the integrative approach in general. At the same time, there are negative implications for tests designed to measure separate linguistic components and skills; for if language proficiency is indeed an inseparable, holistic phenomenon (as this and other research indicates) whose "...parts interact to produce properties and qualities that do not exist in the parts separately," then it follows that a discrete point or separate skills approach must necessarily involve a distorted and inaccurate conception of how this phenomenon actually exists and functions.

In addition to providing a more accurate, more theoretically valid assessment of language proficiency, pragmatic tasks have important advantages in terms of practicality: Not only does a task like dictation, for example, furnish the same information concerning specific abilities (such as spelling, capitalization, and punctuation) as do the tests designed to measure those abilities separately, but because it requires one to "process...sequences of...language constrained by the normal meaningful relationships of...discourse" and to "relate... [them] to extralinguistic context via...pragmatic mappings," it has the potential to yield much richer information concerning one's ability to deal with language in everyday situations. In general, then, pragmatic tasks are quite efficient: They require little time for administration and scoring (all 204 tests were scored by the researcher in less than 50 hours), yet provide accurate and indeed valuable information concerning a student's overall linguistic strengths and weaknesses.

Consequently, if one concludes that not only the two pragmatic subtests, but also the ITBS vocabulary/reading and language skills subtests are all measuring essentially the same construct--global proficiency in the so-called "standard" variety of English--then what would be the implications for a test purporting to measure abstract intelligence if it were

found to intercorrelate as highly (or even higher) with these subtests as it does with other supposed tests of intelligence? This is precisely what we have discovered in this study: The Lorge-Thorndike verbal and nonverbal subtests have been found to intercorrelate quite highly with all four of these subtests. In terms of total scores (for the 7th grade), "Language" was seen to correlate with the ITBS composite at .81 and with the Lorge-Thorndike total at .82, while the latter two intercorrelated at a level approaching the theoretical ideal (.87). In addition, a factor analysis revealed that essentially all of the meaningful variance of the Lorge-Thorndike, the ITBS (excepting arithmetic skills), and the pragmatic tasks was exhausted on a single unitary factor.

In light of these results, it seems unreasonable to assume that the Lorge-Thorndike is any more a measure of abstract intelligence than the ITBS or the cloze and dictation tasks. The verbal subtest--in requiring an individual to deal with words totally void of any extralinguistic context--is more reasonably viewed as a discrete point language test which (like the ITBS language skills and vocabulary/reading subtests and pragmatic tasks) measures global language proficiency much more accurately than it does abstract intelligence. The non-verbal subtest--with its heavy reliance on verbal instructions--also taps this global proficiency to a very large extent, as evidenced by substantial intercorrelations with dictation (.70), ITBS vocabulary/reading (.66), and language skills (.72). Consequently, since all of these subtests exhaust by far the major portion of their variance on a single unitary factor, this factor is most realistically interpreted as "global language proficiency". However, because the non-verbal subtest also intercorrelates substantially with and thus loads on the same factor as the ITBS arithmetic skills subtest (which for the 7th grade clearly taps something other than global proficiency), one must also acknowledge, as did Vineyard and Bailey, that "intelligence tests measure something more than...verbal skills (in this instance quantitative abilities)" (p. 178).

By interpreting the results in such a way, the author is <u>not</u> asserting that (abstract) intelligence does not exist; for indeed it clearly does exist. Instead, what <u>is</u> being asserted here is that because the measures designed to assess intelligence are so greatly dependent upon proficiency in the "standard" variety of English--a proficiency which is affected by numerous factors of which intelligence is but one--they consequently offer a much more accurate assessment of the level of of that proficiency than of abstract intelligence. This is the reason that the designers of the Lorge-Thorndike are able to boast that their test(s) "are excellent predictors of academic success" (Manual, p. 40), for this success as well as performance on tests designed to measure this success--achievement

tests--are both almost entirely dependent upon proficiency in "standard" English. Therefore, the reason that cloze and dictation tasks so accurately predict scores on both the Lorge-Thorndike Intelligence Tests and the Iowa Tests of Basic Skills is that all are essentially measuring the same thing--global language proficiency.

APPENDIX

Name: _____ Grade: _____

PART ONE

Directions: You will hear three short paragraphs. The person on the tape will read each of these paragraphs three times. The first time he reads it, just listen to what he says. The second time he reads the paragraph, write down exactly what he says. He will stop at certain places to give you time to write. The third time he reads it, check your paper to see if you have written exactly what he has said. If you have made a mistake, erase it and write in the correct word or words.

Example: (Listen as he reads it the first time; write as he reads it the second time; check what you have written as he reads it the last time.

Remember: Try to write exactly what the person on the tape says. Be sure just to listen the first time, write the second time, and check what you have written the third time.

(Answer to the example.)

I have two books. One book is about trains. The other book is about dogs.

DICTATIONS
(with pauses shown)

Dictation A

There is / a big tree / near our school. / It is / about one hundred / years old. / During the summer, / the tree has / large green leaves. / These help / to keep the sun / out of my eyes. / In the fall, / the leaves turn / many different colors. / The big tree doesn't have / any leaves at all / during the winter.
Length-55 words. Average length of utterance (ALU)-3.0 words.

Dictation B

Andy has a brand new bike. / The bike is red and blue. / It has a big seat / and a loud horn. / Andy usually rides his bike / to school everyday. / He didn't ride it last Monday / because it rained all day long. / His mother had to take him / to school in the car. / The next day was sunny, / so he rode his bike / all around the neighborhood.
Length-66 words. ALU-5.0 words.

Dictation C

Yesterday I saw a lady / who was walking down a street near my house. / She looked confused and a little bit lost, / so I asked her if she knew where she was going. / She said that she was looking

for Maple Street. / I told her that I would walk with her to Maple Street, / since it was just a couple of blocks away. / I showed her where the street was and then I walked back home.
Length-76 words. ALU-9.5 words.

Dictation D

The stars may or may not have a real effect upon us, / but these days astrology certainly does. / Young people delay their weddings until the zodiac is favorable. / Some companies have astrologers who are constantly kept busy. / Most newspapers print horoscopes and about forty million people can read them. / Even skeptics can't be sure that they are not affected. / Now believers and skeptics alike can read what the stars have to say. / They can also decide for themselves if astrology is really science, fiction, or science fiction.
Length-68 words. ALU-10.8 words.

PART TWO

Directions: You will read three passages. In each passage, there are blanks where a word has been left out. Fill in each blank with the word that you think is missing. You should write only one word in each blank.
Example: I wanted to make myself a sandwich. I walked to the store to _____some bread. When I got to _____store, I realized that I didn't_____ enough money.
After you have read the entire sample, go back and fill in each of the three blanks with the word that you think has been left out. If you change your mind, erase the old word and write in the new one.
(Answer to the example.)
I wanted to make myself a sandwich. I walked to the store to buy some bread. When I got to the store, I realized that I didn't have enough money.
Remember: Only one word in each blank.
Important: You may be able to think of more than one word that can fit correctly in a blank. For example, "I walked to the store to get some bread" is also correct.

CLOZE TESTS

Cloze A

John and Sue live in the same neighborhood. They live in a large city. John has a cat. Its name (1)_____Traveler. John gave his cat this (2)_____ because it likes to travel around (3)_____ neighborhood. Sometimes Traveler goes away and (4)_____ come back all day long. John (5)_____ Sue are standing and talking in (6) _____of John's house.
"Where is your (7)_____, John?" asked Sue. "Here, kitty, kitty, (8)_____."
"I really don't know where he (9) _____," said John. "This makes the second (10) _____this week that he's run away.

(11) _____ just wish he would stay at (12) _____."

"Well, let's walk over to Bill's (13) _____ and look for him there," said (14)_____. "Sometimes he crawls under the porch (15)_____ stays there all afternoon."

So they (16) _____over to Bill's house and asked (17) _____if he had seen Traveler.

"I (18)_____seen him all day long," said (19)_____.

Sue said, "Let's look under your (20)_____and see if he's there."

They bent down and looked under the porch. There was Traveler sound asleep.

Cloze B

Randy and Tom are having a party on Saturday night. They have invited all of their friends. It is the day before the party and they are having some problems.

"Maybe (21)_____shouldn't have invited all of our (22) _____," said Randy. "I only have three (23)_____ and you don't have any money (24)_____all."

"Take it easy, Randy," said (25)_____. "We'll think of something before tomorrow."

"(26)_____ haven't even decided what games we're (27) _____ to have for entertainment," replied Randy. "(28) _____ we don't come up with some (29) _____, we might have to cancel the (30)_____."

"I know what we can do," (31)_____Tom. "Didn't you tell me that (32)_____can beat anyone at the game (33)_____coin tossing?"

"Yes," said Randy. "I've (34)_____tossing coins for five years now."

"(35)_____solves all our problems," said Tom. "(36) _____ don't we have a coin toss (37)_____entertainment?"

"Great! And I'll win enough (38)_____to pay for all the food (39) _____ drinks," exclaimed Randy.

"Isn't that a (40) _____ idea?" asked Tom.

"That's the best idea I've heard in a long time," said Randy. "It's going to be a great party."

Cloze C

I got up early that morning and went out for a little walk. I think it must have been 7 a.m. It was the second week of January and the temperature was only about 20 degrees above zero. I had to (41) _____ on a warm sweater and even (42) _____ heavy jacket because it was so (43) _____. I wore gloves and I even (44)_____to put a scarf around my (45)_____to keep warm and comfortable. The (46)_____ was fair; I could see that (47) _____ was going to be a nice (48)_____.

I was surprised that there was (49) _____ outside but me. That seemed quite (50)_____; after all, 7 a.m. isn't really very (51)_____. I asked myself why nobody was (52)_____ the street: Could my watch be (53)_____ ? Was it really only

5 a.m., and (54) _____ seven? I really didn't know.

After (55) _____ another block without meeting anyone, I (56)_____ a newsboy who was delivering papers (57)_____ his bicycle. "Why are the papers (58) _____ thick today?" I wondered. Like a (59)_____ of lightening, the reason quickly flashed (60)_____ my head: It was Sunday!!!

Cloze D

It was a night that Jesse Richards would remember for many years to come. He had gone to (61)_____ early that evening and it wasn't (62)_____ before he fell asleep. He dreamed (63)_____ he was flying away from the (64)_____ in a huge rocketship. Jesse could (65)_____ the blue sky and the world (66)_____ him. The earth was like a (67)_____ and white ball, getting smaller and (68)_____ with the passing of each second.

(69)_____ a short time, Jesse reached a (70)_____ whose name he never would remember. (71)_____ planet was beautiful; it provided a (72)_____ for thousands of people whose ancestors (73) _____ lived there for many centuries. These (74) _____ were highly skilled at building machines: (75)_____ gigantic telescopes they viewed the entire (76) _____. They had no telephones. Instead, they (77)_____ with each other by television. Jesse (78)_____ knew what the people were saying, (79)_____ the language they spoke was impossible (80) _____ him to understand.

Suddenly, Jesse woke up and heard the sound of the television. He jumped out of bed and ran downstairs. There was his father watching "Star Trek"!!!